UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

Nunez, et al.

v.

The City of New York, et al.

Case No. 11-cv-5845 (LTS)

DECLARATION OF LEIGHANN STARKEY

1.   I am the Director of Data at the Legal Aid Society in New York, New York. I submit this declaration in support of Plaintiffs' reply, filed on February 15, 2023.

2.   I am an experienced data analyst and quantitative researcher, possessing a PhD in the social sciences and having published quantitative research in multiple peer-reviewed, scholarly journals. A true and accurate copy of my curriculum vitae is attached.

3.   I have acted as the Director of Data at the Legal Aid Society since August 2020, where I supervise a team of social science researchers who analyze data relevant to both individual cases and larger advocacy issues.

4.   I have been asked by Plaintiffs' counsel to assess the methodology used by the New York City Department of Correction to audit the accuracy of their New Admissions Dashboard ("Dashboard") in tracking the length of time that individuals who are newly admitted to Department of Correction custody spend in the intake process. In completing this assessment, I have reviewed the following information: the description of audit methodology as described in the February 3, 2023 Monitor report (ECF No. 504, at 18), the representation that "roughly 1,400 people" were processed through intake in January 2023 (*id*. at 17), and the representation that there have been two audits conducted examining sixteen total people (*id*. at 21-22). I am also

informed that the audits resulted in a finding that for 10 of the 16 people, or 63%, housing time entries were accurate. *Id.*

5.   After reviewing these materials, my opinion is that the Department of Correction's current audit methodology is fundamentally flawed, and therefore unable to offer any meaningful assessments of the Dashboard's reliability.

6.   While the level of precision needed in this context is not as high as it might be in some other contexts, such as academic research, there are still basic principles that must not be violated to ensure that samples provide useful and representative information. The practices that the Department of Correction is using are so misguided that their audit results cannot be used to draw any conclusion about the accuracy of the Dashboard as a whole. In other words, the current audit process falls short not only of rigorous academic standards, but also of the minimum standards needed for *any* useful audit.

7.   First, the sampling methodology itself is flawed because it does not represent a random or representative selection process. The current methodology is reportedly to "select a specific day/time from the previous week and to audit the data entered in the New Admissions Dashboard associated with each newly admitted person observed in intake during the selected time period." ECF No. 504, at 18. Selecting a sample using specific days and times is not representative of the conditions that likely impact data entry accuracy. The results of such an audit will inevitably be conflated with specific shifts, personnel working those shifts, and other factors that may change depending on the time of day, day of the week, or the people entering data. A particular day or shift could represent the most reliable data-entry staff, or the least amount of chaos interfering with data entry. A reliable audit sample must account for variation that occurs across day, time, or person. Additionally, it is not indicated that the selection of

specific cases occurred at random. Cases must first be selected such that they are representative of data-entry conditions, and then selected blindly and at-random.

8.   Even if the audit represented a truly random sample, which it does not, the small size of the sample from January 2023 (16 people) relative to the total population being assessed (roughly 1,400 people in that month) creates a margin of error so large that it renders the audit inadequate to meaningfully assess reliability.

9.   The margin of error is the possible difference between the results gleaned from a sample of a population and the results that would be gleaned from the total population. The larger percentage of a total population a random sample represents, the smaller the margin of error. The margin of error is expressed as percentage +/-.

10.   A truly random and representative sample of 16 people out of a total population of 1,400 equates to a +/- 25% margin of error (at a 95% confidence level and assuming an even distribution of data entry errors). This means that even if such an audit of 16 people out of 1400 found 100% of entries to be accurate, the true rate of accuracy for the broader population could be as low as 75%. This margin of error is therefore too large to generalize the 100% accuracy rate to the rest of the population.

11.   The audit results reported thus far for January 2023 found 63% of housing time entries were accurate. ECF No. 504, at 21. Even if these audits used a truly random and representative sample, which they did not, the true percentage of accurate housing time entries for the entire population *could be as low as 38%*.

12.   Even if the audit doubled in size, randomly and representatively sampling 32 of 1,400 people, the margin of error would only decrease to 17%. This would mean that if an audit

of that size reported 63% accuracy, in reality the population-wide accuracy of Dashboard entries could be as low as 46%.
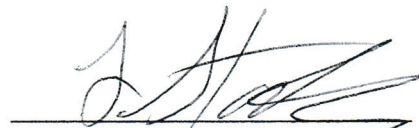
13.   Due to the use of a non-random sampling methodology, as well as the high margin of error created by using such a small sample, in my opinion it is not possible to use the audits conducted thus far to draw any reliable conclusion about the accuracy of the data entered into the Dashboard for the population as a whole.

14.   Furthermore, it is my opinion that conducting additional audits over time using this same methodology will not fix the problem. Because the current process is foundationally flawed, additional audits will continue to provide non-representative results that cannot be used to draw any reliable conclusions about the overall accuracy of the Dashboard.

Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury that the foregoing is true and correct.

February 15, 2023

New York, NY

Leighann Starkey
Director of Data
The Legal Aid Society

Leighann Starkey, Ph.D.

Email: Lnstarkey@gmail.com

Education

2016	Ph.D., Psychology: Human Development, CUNY, Graduate Center, New York, NY
2008	B.S., Psychology, University of Pittsburgh, Pittsburgh, PA

Research and Evaluation

2020-Present

Director of Data, Legal Aid Society

- Direct the production of high-quality analytics and reporting needed to support clients, special litigation projects, policy reform, advocacy efforts, and other objectives within the Criminal Defense Practice.
- Liaise between attorneys, non-attorney staff, and IT to improve organizational data collection and storage practices.
- Procured $40k grant to pilot technical enhancements to current data collection practices.
- Develop and maintain Law Enforcement LookUp, an application dedicated to documenting NYPD misconduct.
- Supervise a team of PhD-level researchers

2017-2020

Director of Evaluation, Education Through Music (ETM)

- Lead and develop the Evaluation Department of ETM, a non-profit organization that partners with New York City schools to provide music as a core subject for all students.
- Served as a member of Senior Leadership, responsible for developing evaluation initiatives to improve programming and demonstrate impact.
- Designed and implemented impact evaluation, including contextualizing ETM's program model in current theories of education; conceptualizing methods and procedures of data collection; gaining IRB approval; overseeing data collection; analyzing and reporting findings to various stakeholders; and making recommendations for program refinement on new and existing program initiatives.

2016-2017

Research Director, National School Climate Center (NSCC)

- Lead the Research Department of NSCC, an internationally-recognized non-profit focusing on school climate measurement and reform.
- Served as a core member of executive leadership, responsible for driving the growth and oversight of large-scale educational research projects with states, districts, and networks nationwide.
- Developed a comprehensive, strategic plan to grow the Research Department, involving implementing a strong equity lens and systems-level approach, streamlining pre-existing projects, identifying and securing funding, and developing new proposals.

2012-2016

Project Director, OPEQ, New York University

- Directed an experimental evaluation of Opportunities for Equitable Access to Quality Basic Education (OPEQ), an intervention aimed at improving teachers' abilities to advance children's math, reading, and socio-emotional skills in the Democratic Republic of Congo (DRC).
- Principal Investigators: Lawrence Aber, PhD, Department of Psychology, NYU and the International Rescue Committee.
- Designed and executed geo-data collection in rural DRC, including training local data collectors.

2011-2012

Policy Research Associate, Institute of Children, Poverty, and Homelessness

2006-2010

Research Assistant, Fordham University and University of Pittsburgh

Teaching

2021-Present
    Adjunct Assistant Professor, CUNY School of Professional Studies (Master's program)
- Perspectives on Developmental Psychology
- Atypical Development

2011- 2013
    Adjunct Lecturer, Brooklyn College (Undergraduate program)
- Children in Crisis
- Introduction to Child Psychology
- Psychological and Developmental Disorders of Childhood

Publications

Starkey, L., Aber, J.L., & Crossman, A. (2019). Risk or resource: Does school climate moderate the influence of community violence on children's social-emotional development in the Democratic Republic of Congo? *Developmental Science*, 22(5)

Torrente, C., Aber, J.L., **Starkey**, L., Johnston, B., Shivshanker, A., Weisenhorn, N., Annan, J., Seidman, E., Wolf, S., & Tubbs Dolan, C. (2019). Improving Primary Education in the Democratic Republic of the Congo: End-Line Results of a Cluster-Randomized Wait-List Controlled Trial of Learning in a Healing Classroom. *Journal of Research on Educational Effectiveness*, DOI: 10.1080/19345747.2018.1561963

Aber, J.L., Tubbs, C., Torrente, C., Halpin, P., Johnston, B., **Starkey, L.,** Shivshanker, A., Annan, J., Seidman, E., & Wolf, S. (2017). Promoting Children's Learning in Conflict-Affected Countries: Testing Change Process in the Democratic Republic of Congo. *Development and Psychopathology,* 29(1): 53-67.

Starkey, L., & Crossman, A. (2017). Emotions in adolescence and gender. In *The SAGE Encyclopedia of Psychology and Gender*. (1st Ed., pp. 494-497). Thousand Oaks: SAGE Publications, Inc.

Starkey, L. (2016). Community violence, school climate, and social-emotional well-being in the Democratic Republic of Congo. *CUNY Academic Works. http://academicworks.cuny.edu/gc_etds/1377*

Aber, J.L., Torrente, C., **Starkey**, L., Johnston, B., Seidman, E., Halpin, P., Shivshanker, A., Weisenhorn, N., Annan, J., & Wolf, S. (2016). Impacts of "Healing Classrooms" on children's reading and math skills in DRC. *Journal of Research on Educational Effectiveness,* 0(0): 1-23.

Torrente, C., Johnston, B., **Starkey**, L., Seidman, E., Shivshanker, A., Weisenhorn, N., Annan, J., & Aber, J.L. (2015). Improving the quality of school interactions and student wellbeing: Impacts after one year of a school-based program in the Democratic Republic of Congo. *Journal on Education in Emergencies*, 1(1): 48-91.

Starkey, L. & Revenson, T.A. (2014). Early changes in socioeconomic status do not predict changes in body mass in the first decade of life. *Annals of Behavioral Medicine*, 49(2): 212-220. DOI: 10.1007/s12160-014-9648-1

Conferences

Starkey, L. (2019, November). Re-thinking "impact" for arts education initiatives: Rejecting test scores and embracing subjectivity. Paper submitted for the Annual Conference of the American Evaluation Association, Minneapolis, MN.

Starkey, L., & Aber, J. (2016, March). Resilience in the Face of Community Violence: Schools as Contexts for Social-Emotional Development in the D.R. Congo. Paper symposium submitted for the Spring 2016 conference for the Society for Research in Adolescence, Baltimore, MD.

Starkey, L., Aber, J., & Johnston, B. (2014, March). Improving Congolese Children's Early Math and Reading Ability: Preliminary Results from a Cluster Randomized Trial in the Democratic Republic of the Congo. Paper submitted for the Spring 2014 conference of the Society for Research on Educational Effectiveness, Washington, D.C.

Starkey, L., & Revenson, T.A. (2014, March). Impact of socioeconomic mobility during early childhood: do early changes in SES predict changes in BMI? Paper submitted for the Biennial Conference of the Society for Research in Human Development, Austin, TX.

Starkey, L. (2013, April). Impact of Neighborhood Food Retail on Internalization of Early Adolescent Nutritional Practices. In G. Evans (Chair), *SES, Child Health, and Well-Being across Contexts and Developmental Stages*. Poster symposium submitted for the Biennial Meeting of the Society for Research in Child Development, Seattle, WA.
.

Software
   Highly proficient in R, STATA, SPSS, MPLUS, Qualtrics, Microsoft Office Suite, Airtable, and SQL.